GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: December 6, 2002, 16:39:25 ; Search time 231.5 Seconds
(without alignments)
14834.978 Million cell updates/sec

Title: US-10-025-514-15
Perfect score: 1525
Sequence: 1 tctagaccatggaagaccct.....ccagtcaaggcctagtcgac 1525

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1525	100.0	1525	24	ABK88025	DNA sequence encod
2	1197.4	78.5	1582	24	ABK88027	DNA sequence encod
3	1197.4	78.5	1756	24	ABK88026	DNA sequence encod
4	1191.6	78.1	1525	24	ABK88022	DNA sequence encod
5	1188.8	78.0	1756	24	ABK88023	DNA sequence encod
6	1187	77.8	1582	24	ABK88024	DNA sequence encod
7	1182	77.5	1182	24	ABK88015	DNA encoding human
8	628.4	41.2	1260	19	AAV41730	Codon-optimised RA
9	432.2	28.3	1312	16	AAQ89254	Human alpha-1-tryp
10	432.2	28.3	1312	19	AAV28471	Nucleotide sequenc
11	432.2	28.3	1312	21	AAZ90199	Human alpha1-antit
12	429.2	28.1	1367	22	AAS45052	cDNA encoding nove
13	429	28.1	1185	19	AAV41726	Native coding sequ
14	429	28.1	1352	13	AAQ31403	Human alpha-1 anti
15	429	28.1	1352	24	ABL67511	Thyroid cancer rel
16	429	28.1	1371	24	ABK84495	Human cDNA differe
17	429	28.1	1371	24	ABL67510	Thyroid cancer rel
18	429	28.1	1433	10	AAN91077	Sequence encoding
19	429	28.1	1434	5	AAN40078	Sequence encoding
20	429	28.1	1434	20	AAX83548	Human alpha1-anti-
21	429	28.1	5932	21	AAZ45928	Nucleotide sequenc
22	429	28.1	6142	21	AAZ45932	Nucleotide sequenc
23	429	28.1	6142	21	AAZ45933	Nucleotide sequenc
24	429	28.1	6565	21	AAZ45925	Nucleotide sequenc
25	429	28.1	6714	21	AAZ45930	Nucleotide sequenc
26	429	28.1	6924	21	AAZ45934	Nucleotide sequenc
27	429	28.1	6924	21	AAZ45935	Nucleotide sequenc
28	429	28.1	6981	21	AAZ45931	Nucleotide sequenc
29	429	28.1	7054	21	AAZ45927	Nucleotide sequenc
30	428.6	28.1	7405	21	AAZ45926	Nucleotide sequenc
31	427.6	28.0	1189	13	AAQ21125	Alpha-1-antitrypsi
32	427.4	28.0	1352	18	AAT72858	Human alpha-1-anti
33	425.8	27.9	1312	10	AAN97127	Sequence of alpha-
34	425.8	27.9	1434	10	AAN90341	Sequence of alpha-
35	425.4	27.9	1185	7	AAN60417	Human alpha 1-anti
36	424.2	27.8	1378	13	AAQ23746	Alpha-1 antitrypsi
37	424.2	27.8	1396	11	AAQ03184	Entire sequence of
38	422.6	27.7	1423	6	AAN50425	Sequence encoding
39	421	27.6	1299	6	AAN50540	Sequence of human
40	421	27.6	1378	6	AAN50021	Sequence encoding
41	407	26.7	1390	22	AAH23089	Osteoarthritis tis
42	402.6	26.4	2013	24	ABL59152	Sequence of fusion
43	372.6	24.4	1242	18	AAT79493	Protease inhibitor
44	359.8	23.6	1242	18	AAT78180	Recombinant squirr
45	357	23.4	1312	10	AAN91078	Alpha-1-antitrypsi

ALIGNMENTS

RESULT 1
ABK88025
ID ABK88025 standard; DNA; 1525 BP.
XX
AC ABK88025;
XX
DT 07-OCT-2002 (first entry)
XX
DE DNA sequence encoding rSLAP1 fusion protein.
XX
KW rSLAP1; gene; ds; Alzheimer's disease; tumour angiogenesis;
KW malaria; emphysema; asthma; chronic obstructive pulmonary disease;
KW cystic fibrosis; otitis media; otitis external; HIV; psoriasis; eczema;
KW human immunodeficiency virus; atopic dermatitis; muscular dystrophy;
KW herpes; ulceration; sepsis; rheumatoid arthritis; periodontal disease;
KW tumour metastasis; osteoporosis; Paget's disease; scleroderma;
KW glomerulonephritis; hypertension.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT RBS 6..8
FT /*tag= a
FT /standard_name= "Ribosome binding site"
FT CDS 9..1520
FT /*tag= b
FT /product= "rSLAP1 fusion protein"
FT misc_feature 12..1193
FT /*tag= c
FT /note= "AAT coding region"
FT misc_feature 1194..1196

```
FT                   /*tag=  d
FT                   /note= "linking codon"
FT   misc_feature    1197..1517
FT                   /*tag=  e
FT                   /note= "SLPI coding region"
XX
PN   WO200250287-A2.
XX
PD   27-JUN-2002.
XX
PF   18-DEC-2001; 2001WO-US49256.
XX
PR   18-DEC-2000; 2000US-256699P.
PR   20-NOV-2001; 2001US-331966P.
XX
PA   (ARRI-) ARRIVA PHARM INC.
XX
PI   Barr PJ,  Gibson HL,  Pemberton P;
XX
DR   WPI; 2002-500631/53.
DR   P-PSDB; AAU99884.
XX
PT   Novel fusion protein useful for inhibiting protease activity associated
PT   with a disorder such as emphysema, asthma, comprises a first protease
PT   inhibitor comprising alpha 1-antitrypsin and a second protease
PT   inhibitor -
XX
PS   Example 3; Page 89-90; 134pp; English.
XX
CC   This invention relates to a novel fusion protein comprising a first
CC   protease inhibitor comprising an alphal-antitrypsin or its functionally
CC   active portion and a second protease inhibitor or its functionally
CC   active protein. The fusion proteins of the invention may act as an
CC   inhibitor of protease activity. The fusion protein of the invention
CC   is useful for inhibiting protease activity associated with a disorder
CC   such as emphysema, asthma, chronic obstructive pulmonary disease,
CC   cystic fibrosis, otitis media, otitis external or HIV infection, or
CC   for treating an individual suffering from or at risk for a disease or
CC   disorder involving unwanted protease activity. The proteins are useful
CC   for treating dermatological diseases such as atopic dermatitis, eczema
CC   and psoriasis, in inflammatory responses to viral infection, and for
CC   treating herpes infection, corneal or epidermal ulceration, chronic
CC   non-healing wounds, sepsis, rheumatoid arthritis, periodontal disease,
CC   tumour metastasis and tumour angiogenesis, gastric ulceration,
CC   osteoporosis, Paget's disease, glomerulonephritis, scleroderma, malaria,
CC   bacterial infection, Alzheimer's disease, hypertension and muscular
CC   dystrophy. The present sequence represents the DNA encoding the
CC   rSLAP1 fusion protein of the invention.
XX
SQ   Sequence 1525 BP; 467 A; 287 C; 314 G; 457 T; 0 other;
```

Query Match 100.0%; Score 1525; DB 24; Length 1525;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1525; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 TCTAGACCATGGAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACG 60
Db       1 TCTAGACCATGGAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACG 60

Qy      61 ACCAAGACCATCCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTT 120
Db      61 ACCAAGACCATCCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTT 120

Qy     121 TGTATAGACAATTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTA 180
Db     121 TGTATAGACAATTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTA 180

Qy     181 TTGCCACTGCTTTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTT 240
Db     181 TTGCCACTGCTTTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTT 240

Qy     241 TAGAAGGTTTAAACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTC 300
Db     241 TAGAAGGTTTAAACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTC 300

Qy     301 AAGAGTTGTTGAGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACG 360
Db     301 AAGAGTTGTTGAGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACG 360

Qy     361 GTTTATTTTTGTCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAAC 420
Db     361 GTTTATTTTTGTCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAAC 420

Qy     421 TATATCATAGTGAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAA 480
Db     421 TATATCATAGTGAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAA 480

Qy     481 TTAATGATTATGTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAG 540
Db     481 TTAATGATTATGTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAG 540

Qy     541 ATCGTGATACCGTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTC 600
Db     541 ATCGTGATACCGTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTC 600

Qy     601 CTTTCGAGGTTAAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCA 660
Db     601 CTTTCGAGGTTAAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCA 660

Qy     661 AAGTTCCAATGATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTT 720
Db     661 AAGTTCCAATGATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTT 720

Qy     721 CTTGGGTCTTATTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACG 780
Db     721 CTTGGGTCTTATTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACG 780

Qy     781 AAGGTAAGCTTCAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAG 840
Db     781 AAGGTAAGCTTCAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAG 840

Qy     841 AGAACGAGGATCGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTT 900
Db     841 AGAACGAGGATCGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTT 900

Qy     901 ACGACTTAAAATCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCG 960
Db     901 ACGACTTAAAATCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCG 960

Qy     961 ATTTGAGTGGTGTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCG 1020
Db     961 ATTTGAGTGGTGTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCG 1020

Qy    1021 TCTTAACTATTGATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTC 1080
Db    1021 TCTTAACTATTGATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTC 1080

Qy    1081 CAATGAGCATTCCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGC 1140
Db    1081 CAATGAGCATTCCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGC 1140

Qy    1141 AGAACACTAAAAGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAAGATGTCCG 1200
Db    1141 AGAACACTAAAAGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAAGATGTCCG 1200

Qy    1201 GAAAGTCTTTCAAGGCCGGTGTTTGTCCACCAAAGAAGTCCGCTCAATGTTTGAGATACA 1260
Db    1201 GAAAGTCTTTCAAGGCCGGTGTTTGTCCACCAAAGAAGTCCGCTCAATGTTTGAGATACA 1260

Qy    1261 AGAAGCCAGAATGTCAATCCGACTGGCAATGTCCAGGTAAGAAGAGATGTTGTCCAGACA 1320
Db    1261 AGAAGCCAGAATGTCAATCCGACTGGCAATGTCCAGGTAAGAAGAGATGTTGTCCAGACA 1320

Qy    1321 CTTGTGGTATCAAGTGTCTAGACCCAGTTGACACCCCAAACCCAACTAGAAGAAAGCCAG 1380
Db    1321 CTTGTGGTATCAAGTGTCTAGACCCAGTTGACACCCCAAACCCAACTAGAAGAAAGCCAG 1380
```

```
Qy   1381 GTAAGTGTCCAGTTACTTACGGTCAATGTTTGATGTTGAACCCACCAAACTTCTGTGAAA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GTAAGTGTCCAGTTACTTACGGTCAATGTTTGATGTTGAACCCACCAAACTTCTGTGAAA 1440

Qy   1441 TGGACGGTCAATGTAAGAGAGACTTGAAGTGTTGTATGGGTATGTGTGGTAAGTCCTGTG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TGGACGGTCAATGTAAGAGAGACTTGAAGTGTTGTATGGGTATGTGTGGTAAGTCCTGTG 1500

Qy   1501 TTTCCCCAGTCAAGGCCTAGTCGAC 1525
          |||||||||||||||||||||||||
Db   1501 TTTCCCCAGTCAAGGCCTAGTCGAC 1525
```

RESULT 2
ABK88027
ID ABK88027 standard; DNA; 1582 BP.
XX
AC ABK88027;
XX
DT 07-OCT-2002 (first entry)
XX
DE DNA sequence encoding rN-TAP1 fusion protein.
XX
KW rN-TAP1; gene; ds; Alzheimer's disease; tumour angiogenesis;
KW malaria; emphysema; asthma; chronic obstructive pulmonary disease;
KW cystic fibrosis; otitis media; otitis external; HIV; psoriasis; eczema;
KW human immunodeficiency virus; atopic dermatitis; muscular dystrophy;
KW herpes; ulceration; sepsis; rheumatoid arthritis; periodontal disease;
KW tumour metastasis; osteoporosis; Paget's disease; scleroderma;
KW glomerulonephritis; hypertension.
XX
OS Homo sapiens.
OS Synthetic.
XX

FH	Key	Location/Qualifiers
FT	RBS	6..8
FT		/*tag= a
FT		/standard_name= "Ribosome binding site"
FT	CDS	9..1577
FT		/*tag= b
FT		/product= "rTAP1 fusion protein"
FT	misc_feature	12..1193
FT		/*tag= c
FT		/note= "AAT coding region"
FT	misc_feature	1194..1196
FT		/*tag= d
FT		/note= "linking codon"
FT	misc_feature	1197..1574
FT		/*tag= e
FT		/note= "TIMP-1 coding region"

XX
PN WO200250287-A2.
XX
PD 27-JUN-2002.
XX
PF 18-DEC-2001; 2001WO-US49256.
XX
PR 18-DEC-2000; 2000US-256699P.
PR 20-NOV-2001; 2001US-331966P.
XX
PA (ARRI-) ARRIVA PHARM INC.
XX
PI Barr PJ, Gibson HL, Pemberton P;
XX
DR WPI; 2002-500631/53.
DR P-PSDB; AAU99885.
XX
PT Novel fusion protein useful for inhibiting protease activity associated
PT with a disorder such as emphysema, asthma, comprises a first protease
PT inhibitor comprising alpha 1-antitrypsin and a second protease
PT inhibitor -
XX
PS Example 3; Page 95-96; 134pp; English.
XX
CC This invention relates to a novel fusion protein comprising a first
CC protease inhibitor comprising an alpha1-antitrypsin or its functionally
CC active portion and a second protease inhibitor or its functionally
CC active protein. The fusion proteins of the invention may act as an
CC inhibitor of protease activity. The fusion protein of the invention
CC is useful for inhibiting protease activity associated with a disorder
CC such as emphysema, asthma, chronic obstructive pulmonary disease,
CC cystic fibrosis, otitis media, otitis external or HIV infection, or
CC for treating an individual suffering from or at risk for a disease or
CC disorder involving unwanted protease activity. The proteins are useful
CC for treating dermatological diseases such as atopic dermatitis, eczema
CC and psoriasis, in inflammatory responses to viral infection, and for
CC treating herpes infection, corneal or epidermal ulceration, chronic
CC non-healing wounds, sepsis, rheumatoid arthritis, periodontal disease,
CC tumour metastasis and tumour angiogenesis, gastric ulceration,
CC osteoporosis, Paget's disease, glomerulonephritis, scleroderma, malaria,
CC bacterial infection, Alzheimer's disease, hypertension and muscular
CC dystrophy. The present sequence represents the DNA encoding the
CC rN-TAP1 fusion protein of the invention.
XX
SQ Sequence 1582 BP; 464 A; 334 C; 329 G; 455 T; 0 other;

Query Match 78.5%; Score 1197.4; DB 24; Length 1582;
Best Local Similarity 99.9%; Pred. No. 5.5e-291;
Matches 1198; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 TCTAGACCATGGAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TCTAGACCATGGAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACG 60

Qy     61 ACCAAGACCATCCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACCAAGACCATCCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTT 120

Qy    121 TGTATAGACAATTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGTATAGACAATTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTA 180

Qy    181 TTGCCACTGCTTTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TTGCCACTGCTTTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTT 240

Qy    241 TAGAAGGTTTAAACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TAGAAGGTTTAAACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTC 300

Qy    301 AAGAGTTGTTGAGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AAGAGTTGTTGAGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACG 360

Qy    361 GTTTATTTTTGTCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GTTTATTTTTGTCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAAC 420

Qy    421 TATATCATAGTGAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TATATCATAGTGAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAA 480

Qy    481 TTAATGATTATGTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TTAATGATTATGTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAG 540

Qy    541 ATCGTGATACCGTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ATCGTGATACCGTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTC 600

Qy    601 CTTTCGAGGTTAAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTTCGAGGTTAAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCA 660

Qy    661 AAGTTCCAATGATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTT 720
```

```
Db    661 AAGTTCCAATGATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTT 720

Qy    721 CTTGGGTCTTATTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CTTGGGTCTTATTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACG 780

Qy    781 AAGGTAAGCTTCAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AAGGTAAGCTTCAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAG 840

Qy    841 AGAACGAGGATCGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AGAACGAGGATCGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTT 900

Qy    901 ACGACTTAAAATCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ACGACTTAAAATCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCG 960

Qy    961 ATTTGAGTGGTGTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATTTGAGTGGTGTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCG 1020

Qy   1021 TCTTAACTATTGATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TCTTAACTATTGATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTC 1080

Qy   1081 CAATGAGCATTCCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGC 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAATGAGCATTCCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGC 1140

Qy   1141 AGAACACTAAAAGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAAGATGTCC 1199
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   1141 AGAACACTAAAAGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAAGATGTGC 1199
```

```
RESULT 3
ABK88026
ID   ABK88026 standard; DNA; 1756 BP.
XX
AC   ABK88026;
XX
DT   07-OCT-2002  (first entry)
XX
DE   DNA sequence encoding rTAP1 fusion protein.
XX
KW   rTAP1; gene; ds; Alzheimer's disease; tumour angiogenesis;
KW   malaria; emphysema; asthma; chronic obstructive pulmonary disease;
KW   cystic fibrosis; otitis media; otitis external; HIV; psoriasis; eczema;
KW   human immunodeficiency virus; atopic dermatitis; muscular dystrophy;
KW   herpes; ulceration; sepsis; rheumatoid arthritis; periodontal disease;
KW   tumour metastasis; osteoporosis; Paget's disease; scleroderma;
KW   glomerulonephritis; hypertension.
XX
OS    Homo sapiens.
OS    Synthetic.
XX
FH   Key             Location/Qualifiers
FT   RBS             6..8
FT                   /*tag=  a
FT                   /standard_name= "Ribosome binding site"
FT   CDS             9..1751
FT                   /*tag=  b
FT                   /product= "rTAP1 fusion protein"
FT   misc_feature    12..1193
FT                   /*tag=  c
FT                   /note= "AAT coding region"
FT   misc_feature    1194..1196
FT                   /*tag=  d
FT                   /note= "linking codon"
FT   misc_feature    1197..1748
FT                   /*tag=  e
FT                   /note= "TIMP-1 coding region"
```

```
XX
PN   WO200250287-A2.
XX
PD   27-JUN-2002.
XX
PF   18-DEC-2001; 2001WO-US49256.
XX
PR   18-DEC-2000; 2000US-256699P.
PR   20-NOV-2001; 2001US-331966P.
XX
PA   (ARRI-) ARRIVA PHARM INC.
XX
PI   Barr PJ,  Gibson HL,  Pemberton P;
XX
DR   WPI; 2002-500631/53.
DR   P-PSDB; AAU99889.
XX
PT   Novel fusion protein useful for inhibiting protease activity associated
PT   with a disorder such as emphysema, asthma, comprises a first protease
PT   inhibitor comprising alpha 1-antitrypsin and a second protease
PT   inhibitor -
XX
PS   Example 3; Page 92-93; 134pp; English.
XX
CC   This invention relates to a novel fusion protein comprising a first
CC   protease inhibitor comprising an alpha1-antitrypsin or its functionally
CC   active portion and a second protease inhibitor or its functionally
CC   active protein. The fusion proteins of the invention may act as an
CC   inhibitor of protease activity. The fusion protein of the invention
CC   is useful for inhibiting protease activity associated with a disorder
CC   such as emphysema, asthma, chronic obstructive pulmonary disease,
CC   cystic fibrosis, otitis media, otitis external or HIV infection, or
CC   for treating an individual suffering from or at risk for a disease or
CC   disorder involving unwanted protease activity. The proteins are useful
CC   for treating dermatological diseases such as atopic dermatitis, eczema
CC   and psoriasis, in inflammatory responses to viral infection, and for
CC   treating herpes infection, corneal or epidermal ulceration, chronic
CC   non-healing wounds, sepsis, rheumatoid arthritis, periodontal disease,
CC   tumour metastasis and tumour angiogenesis, gastric ulceration,
CC   osteoporosis, Paget's disease, glomerulonephritis, scleroderma, malaria,
CC   bacterial infection, Alzheimer's disease, hypertension and muscular
CC   dystrophy. The present sequence represents the DNA encoding the
CC   rTAP1 fusion protein of the invention.
XX
SQ   Sequence 1756 BP; 493 A; 394 C; 374 G; 495 T; 0 other;
```

```
  Query Match           78.5%;  Score 1197.4;  DB 24;  Length 1756;
  Best Local Similarity 99.9%;  Pred. No. 5.7e-291;
  Matches 1198;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TCTAGACCATGGAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TCTAGACCATGGAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACG 60

Qy     61 ACCAAGACCATCCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACCAAGACCATCCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTT 120

Qy    121 TGTATAGACAATTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGTATAGACAATTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTA 180

Qy    181 TTGCCACTGCTTTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TTGCCACTGCTTTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTT 240

Qy    241 TAGAAGGTTTAAACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TAGAAGGTTTAAACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTC 300

Qy    301 AAGAGTTGTTGAGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AAGAGTTGTTGAGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACG 360
```

```
Qy     361 GTTTATTTTTGTCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAAC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GTTTATTTTTGTCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAAC 420

Qy     421 TATATCATAGTGAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TATATCATAGTGAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAA 480

Qy     481 TTAATGATTATGTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TTAATGATTATGTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAG 540

Qy     541 ATCGTGATACCGTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 ATCGTGATACCGTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTC 600

Qy     601 CTTTCGAGGTTAAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCA 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CTTTCGAGGTTAAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCA 660

Qy     661 AAGTTCCAATGATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTT 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AAGTTCCAATGATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTT 720

Qy     721 CTTGGGTCTTATTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTTACCAGACG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CTTGGGTCTTATTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTTACCAGACG 780

Qy     781 AAGGTAAGCTTCAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 AAGGTAAGCTTCAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAG 840

Qy     841 AGAACGAGGATCGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 AGAACGAGGATCGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTT 900

Qy     901 ACGACTTAAAATCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ACGACTTAAAATCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCG 960

Qy     961 ATTTGAGTGGTGTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ATTTGAGTGGTGTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCG 1020

Qy    1021 TCTTAACTATTGATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 TCTTAACTATTGATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTC 1080

Qy    1081 CAATGAGCATTCCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGC 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CAATGAGCATTCCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGC 1140

Qy    1141 AGAACACTAAAAGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAAGATGTCC 1199
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    1141 AGAACACTAAAAGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAAGATGTGC 1199
```

RESULT 4
ABK88022
ID ABK88022 standard; DNA; 1525 BP.
XX
AC ABK88022;
XX
DT 07-OCT-2002 (first entry)
XX
DE DNA sequence encoding SLAP1 fusion protein.
XX
KW Yeast; alpha factor; gene; ds; Alzheimer's disease; SLAP1;
KW malaria; emphysema; asthma; chronic obstructive pulmonary disease;
KW cystic fibrosis; otitis media; otitis external; HIV; psoriasis; eczema;
KW human immunodeficiency virus; atopic dermatitis; muscular dystrophy;
KW herpes; ulceration; sepsis; rheumatoid arthritis; periodontal disease;
KW tumour metastasis; tumour angiogenesis; osteoporosis; Paget's disease;
KW glomerulonephritis; scleroderma; hypertension.
XX
OS Homo sapiens.
OS Synthetic.
XX

FH	Key	Location/Qualifiers
FT	RBS	6..8
FT		/*tag= a
FT		/standard_name= "Ribosome binding site"
FT	CDS	9..1520
FT		/*tag= b
FT		/product= "SLAP1 fusion protein"
FT	misc_feature	12..332
FT		/*tag= c
FT		/note= "SLPI coding region"
FT	misc_feature	333-335
FT		/*tag= d
FT		/note= "linking codon"
FT	misc_feature	336..1517
FT		/*tag= e
FT		/note= "AAT coding region"

XX
PN WO200250287-A2.
XX
PD 27-JUN-2002.
XX
PF 18-DEC-2001; 2001WO-US49256.
XX
PR 18-DEC-2000; 2000US-256699P.
PR 20-NOV-2001; 2001US-331966P.
XX
PA (ARRI-) ARRIVA PHARM INC.
XX
PI Barr PJ, Gibson HL, Pemberton P;
XX
DR WPI; 2002-500631/53.
DR P-PSDB; AAU99881.
XX
PT Novel fusion protein useful for inhibiting protease activity associated
PT with a disorder such as emphysema, asthma, comprises a first protease
PT inhibitor comprising alpha 1-antitrypsin and a second protease
PT inhibitor -
XX
PS Example 1; Page 73-73; 134pp; English.
XX
CC This invention relates to a novel fusion protein comprising a first
CC protease inhibitor comprising an alpha1-antitrypsin or its functionally
CC active portion and a second protease inhibitor or its functionally
CC active protein. The fusion proteins of the invention may act as an
CC inhibitor of protease activity. The fusion protein of the invention
CC is useful for inhibiting protease activity associated with a disorder
CC such as emphysema, asthma, chronic obstructive pulmonary disease,
CC cystic fibrosis, otitis media, otitis external or HIV infection, or
CC for treating an individual suffering from or at risk for a disease or
CC disorder involving unwanted protease activity. The proteins are useful
CC for treating dermatological diseases such as atopic dermatitis, eczema
CC and psoriasis, in inflammatory responses to viral infection, and for
CC treating herpes infection, corneal or epidermal ulceration, chronic
CC non-healing wounds, sepsis, rheumatoid arthritis, periodontal disease,
CC tumour metastasis and tumour angiogenesis, gastric ulceration,
CC osteoporosis, Paget's disease, glomerulonephritis, scleroderma, malaria,
CC bacterial infection, Alzheimer's disease, hypertension and muscular
CC dystrophy. The present sequence represents the DNA encoding the
CC SLAP1 fusion protein of the invention.
XX
SQ Sequence 1525 BP; 467 A; 286 C; 314 G; 458 T; 0 other;

Query Match 78.1%; Score 1191.6; DB 24; Length 1525;
Best Local Similarity 99.7%; Pred. No. 1.6e-289;
Matches 1194; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 TCTAGACCATGGAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACG 60
```

```
Db    325 TCAAGGCCATGGAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACG 384

Qy     61 ACCAAGACCATCCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTT 120

Db    385 ACCAAGACCATCCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTT 444

Qy    121 TGTATAGACAATTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTA 180

Db    445 TGTATAGACAATTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTA 504

Qy    181 TTGCCACTGCTTTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTT 240

Db    505 TTGCCACTGCTTTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTT 564

Qy    241 TAGAAGGTTTAAACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTC 300

Db    565 TAGAAGGTTTAAACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTC 624

Qy    301 AAGAGTTGTTGAGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACG 360

Db    625 AAGAGTTGTTGAGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACG 684

Qy    361 GTTTATTTTTGTCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAAC 420

Db    685 GTTTATTTTTGTCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAAC 744

Qy    421 TATATCATAGTGAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAA 480

Db    745 TATATCATAGTGAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAA 804

Qy    481 TTAATGATTATGTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAG 540

Db    805 TTAATGATTATGTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAG 864

Qy    541 ATCGTGATACCGTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTC 600

Db    865 ATCGTGATACCGTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTC 924

Qy    601 CTTTCGAGGTTAAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCA 660

Db    925 CTTTCGAGGTTAAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCA 984

Qy    661 AAGTTCCAATGATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTT 720

Db    985 AAGTTCCAATGATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTT 1044

Qy    721 CTTGGGTCTTATTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTTACCAGACG 780

Db   1045 CTTGGGTCTTATTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTTACCAGACG 1104

Qy    781 AAGGTAAGCTTCAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAG 840

Db   1105 AAGGTAAGCTTCAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAG 1164

Qy    841 AGAACGAGGATCGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTT 900

Db   1165 AGAACGAGGATCGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTT 1224

Qy    901 ACGACTTAAAATCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCG 960

Db   1225 ACGACTTAAAATCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCG 1284

Qy    961 ATTTGAGTGGTGTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCG 1020

Db   1285 ATTTGAGTGGTGTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCG 1344

Qy   1021 TCTTAACTATTGATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTC 1080

Db   1345 TCTTAACTATTGATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTC 1404

Qy   1081 CAATGAGCATTCCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGC 1140

Db   1405 CAATGAGCATTCCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGC 1464

Qy   1141 AGAACACTAAAAGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAAGATGTC 1198

Db   1465 AGAACACTAAAAGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAAGTAGTC 1522
```

RESULT 5
ABK88023
ID ABK88023 standard; DNA; 1756 BP.
XX
AC ABK88023;
XX
DT 07-OCT-2002 (first entry)
XX
DE DNA sequence encoding TAP1 fusion protein.
XX
KW TAP1; gene; ds; Alzheimer's disease; tumour angiogenesis;
KW malaria; emphysema; asthma; chronic obstructive pulmonary disease;
KW cystic fibrosis; otitis media; otitis external; HIV; psoriasis; eczema;
KW human immunodeficiency virus; atopic dermatitis; muscular dystrophy;
KW herpes; ulceration; sepsis; rheumatoid arthritis; periodontal disease;
KW tumour metastasis; osteoporosis; Paget's disease; scleroderma;
KW glomerulonephritis; hypertension.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key   Location/Qualifiers
FT RBS   6..8
FT   /*tag= a
FT   /standard_name= "Ribosome binding site"
FT CDS   9..1751
FT   /*tag= b
FT   /product= "TAP1 fusion protein"
FT misc_feature   12..563
FT   /*tag= c
FT   /note= "TIMP-1 coding region"
FT misc_feature   564..566
FT   /*tag= d
FT   /note= "linking codon"
FT misc_feature   567..1748
FT   /*tag= e
FT   /note= "AAT coding region"
XX
PN WO200250287-A2.
XX
PD 27-JUN-2002.
XX
PF 18-DEC-2001; 2001WO-US49256.
XX
PR 18-DEC-2000; 2000US-256699P.
PR 20-NOV-2001; 2001US-331966P.
XX
PA (ARRI-) ARRIVA PHARM INC.
XX
PI Barr PJ, Gibson HL, Pemberton P;
XX
DR WPI; 2002-500631/53.
DR P-PSDB; AAU99882.
XX
PT Novel fusion protein useful for inhibiting protease activity associated
PT with a disorder such as emphysema, asthma, comprises a first protease
PT inhibitor comprising alpha 1-antitrypsin and a second protease
PT inhibitor -
XX
PS Example 1; Page 77-78; 134pp; English.
XX
CC This invention relates to a novel fusion protein comprising a first
CC protease inhibitor comprising an alphal-antitrypsin or its functionally
CC active portion and a second protease inhibitor or its functionally
CC active protein. The fusion proteins of the invention may act as an
CC inhibitor of protease activity. The fusion protein of the invention

```
CC   is useful for inhibiting protease activity associated with a disorder
CC   such as emphysema, asthma, chronic obstructive pulmonary disease,
CC   cystic fibrosis, otitis media, otitis external or HIV infection, or
CC   for treating an individual suffering from or at risk for a disease or
CC   disorder involving unwanted protease activity. The proteins are useful
CC   for treating dermatological diseases such as atopic dermatitis, eczema
CC   and psoriasis, in inflammatory responses to viral infection, and for
CC   treating herpes infection, corneal or epidermal ulceration, chronic
CC   non-healing wounds, sepsis, rheumatoid arthritis, periodontal disease,
CC   tumour metastasis and tumour angiogenesis, gastric ulceration,
CC   osteoporosis, Paget's disease, glomerulonephritis, scleroderma, malaria,
CC   bacterial infection, Alzheimer's disease, hypertension and muscular
CC   dystrophy. The present sequence represents the DNA encoding the
CC   TAP1 fusion protein of the invention.
XX
SQ   Sequence 1756 BP; 493 A; 395 C; 373 G; 495 T; 0 other;

  Query Match             78.0%;  Score 1188.8;  DB 24;  Length 1756;
  Best Local Similarity   99.8%;  Pred. No. 8.3e-289;
  Matches 1190;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       7 CCATGGAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAG 66
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     562 CCATGGAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAG 621

Qy      67 ACCATCCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATA 126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     622 ACCATCCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATA 681

Qy     127 GACAATTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCA 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     682 GACAATTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTTAGTCCTGTTTCTATTGCCA 741

Qy     187 CTGCTTTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAG 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     742 CTGCTTTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAG 801

Qy     247 GTTTAAACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGT 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     802 GTTTAAACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGT 861

Qy     307 TGTTGAGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTAT 366
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     862 TGTTGAGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTAT 921

Qy     367 TTTTGTCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATC 426
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     922 TTTTGTCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATC 981

Qy     427 ATAGTGAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATG 486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     982 ATAGTGAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATG 1041

Qy     487 ATTATGTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTG 546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1042 ATTATGTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTG 1101

Qy     547 ATACCGTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCG 606
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1102 ATACCGTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCG 1161

Qy     607 AGGTTAAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTC 666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1162 AGGTTAAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTC 1221

Qy     667 CAATGATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGG 726
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1222 CAATGATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGG 1281

Qy     727 TCTTATTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTA 786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1282 TCTTATTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTA 1341

Qy     787 AGCTTCAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACG 846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1342 AGCTTCAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACG 1401

Qy     847 AGGATCGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACT 906
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1402 AGGATCGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACT 1461

Qy     907 TAAAATCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGA 966
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1462 TAAAATCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGA 1521

Qy     967 GTGGTGTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAA 1026
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1522 GTGGTGTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAA 1581

Qy    1027 CTATTGATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGA 1086
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1582 CTATTGATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGA 1641

Qy    1087 GCATTCCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACA 1146
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1642 GCATTCCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACA 1701

Qy    1147 CTAAAAGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAAGATGTC 1198
          ||||||||||||||||||||||||||||||||||||||||||||||||||||  |||
Db    1702 CTAAAAGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAAGTAGTC 1753


RESULT 6
ABK88024
ID   ABK88024 standard; DNA; 1582 BP.
XX
AC   ABK88024;
XX
DT   07-OCT-2002  (first entry)
XX
DE   DNA sequence encoding N-TAP1 fusion protein.
XX
KW   NTAP1; gene; ds; Alzheimer's disease; tumour angiogenesis;
KW   malaria; emphysema; asthma; chronic obstructive pulmonary disease;
KW   cystic fibrosis; otitis media; otitis external; HIV; psoriasis; eczema;
KW   human immunodeficiency virus; atopic dermatitis; muscular dystrophy;
KW   herpes; ulceration; sepsis; rheumatoid arthritis; periodontal disease;
KW   tumour metastasis; osteoporosis; Paget's disease; scleroderma;
KW   glomerulonephritis; hypertension.
XX
OS    Homo sapiens.
OS    Synthetic.
XX
FH   Key             Location/Qualifiers
FT   RBS             6..8
FT                   /*tag= a
FT                   /standard_name= "Ribosome binding site"
FT   CDS             9..1577
FT                   /*tag= b
FT                   /product= "NTAP1 fusion protein"
FT   misc_feature    12..389
FT                   /*tag= c
FT                   /note= "TIMP-1 coding region"
FT   misc_feature    390..392
FT                   /*tag= d
FT                   /note= "linking codon"
FT   misc_feature    393..1574
FT                   /*tag= e
FT                   /note= "AAT coding region"
XX
PN   WO200250287-A2.
XX
PD   27-JUN-2002.
XX
PF   18-DEC-2001; 2001WO-US49256.
XX
```

PR 18-DEC-2000; 2000US-256699P.
PR 20-NOV-2001; 2001US-331966P.
XX
PA (ARRI-) ARRIVA PHARM INC.
XX
PI Barr PJ, Gibson HL, Pemberton P;
XX
DR WPI; 2002-500631/53.
DR P-PSDB; AAU99883.
XX
PT Novel fusion protein useful for inhibiting protease activity associated
PT with a disorder such as emphysema, asthma, comprises a first protease
PT inhibitor comprising alpha 1-antitrypsin and a second protease
PT inhibitor -
XX
PS Example 2; Page 85-86; 134pp; English.
XX
CC This invention relates to a novel fusion protein comprising a first
CC protease inhibitor comprising an alpha1-antitrypsin or its functionally
CC active portion and a second protease inhibitor or its functionally
CC active protein. The fusion proteins of the invention may act as an
CC inhibitor of protease activity. The fusion protein of the invention
CC is useful for inhibiting protease activity associated with a disorder
CC such as emphysema, asthma, chronic obstructive pulmonary disease,
CC cystic fibrosis, otitis media, otitis external or HIV infection, or
CC for treating an individual suffering from or at risk for a disease or
CC disorder involving unwanted protease activity. The proteins are useful
CC for treating dermatological diseases such as atopic dermatitis, eczema
CC and psoriasis, in inflammatory responses to viral infection, and for
CC treating herpes infection, corneal or epidermal ulceration, chronic
CC non-healing wounds, sepsis, rheumatoid arthritis, periodontal disease,
CC tumour metastasis and tumour angiogenesis, gastric ulceration,
CC osteoporosis, Paget's disease, glomerulonephritis, scleroderma, malaria,
CC bacterial infection, Alzheimer's disease, hypertension and muscular
CC dystrophy. The present sequence represents the DNA encoding the
CC NTAP1 fusion protein of the invention.
XX
SQ Sequence 1582 BP; 464 A; 333 C; 329 G; 456 T; 0 other;

Query Match 77.8%; Score 1187; DB 24; Length 1582;
Best Local Similarity 99.6%; Pred. No. 2.3e-288;
Matches 1190; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy      4 AGACCATGGAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACC 63
          ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    385 AGGAAATGGAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACC 444

Qy     64 AAGACCATCCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGT 123
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    445 AAGACCATCCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGT 504

Qy    124 ATAGACAATTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTG 183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    505 ATAGACAATTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTG 564

Qy    184 CCACTGCTTTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAG 243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    565 CCACTGCTTTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAG 624

Qy    244 AAGGTTTAAACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAG 303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    625 AAGGTTTAAACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAG 684

Qy    304 AGTTGTTGAGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTT 363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    685 AGTTGTTGAGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTT 744

Qy    364 TATTTTTGTCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTAT 423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    745 TATTTTTGTCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTAT 804

Qy    424 ATCATAGTGAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTA 483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    805 ATCATAGTGAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTA 864

Qy    484 ATGATTATGTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATC 543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    865 ATGATTATGTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATC 924

Qy    544 GTGATACCGTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTT 603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    925 GTGATACCGTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTT 984

Qy    604 TCGAGGTTAAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAG 663
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    985 TCGAGGTTAAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAG 1044

Qy    664 TTCCAATGATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTT 723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1045 TTCCAATGATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTT 1104

Qy    724 GGGTCTTATTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTTACCAGACGAAG 783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1105 GGGTCTTATTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTTACCAGACGAAG 1164

Qy    784 GTAAGCTTCAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGA 843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1165 GTAAGCTTCAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGA 1224

Qy    844 ACGAGGATCGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACG 903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1225 ACGAGGATCGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACG 1284

Qy    904 ACTTAAAATCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATT 963
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1285 ACTTAAAATCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATT 1344

Qy    964 TGAGTGGTGTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCT 1023
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1345 TGAGTGGTGTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCT 1404

Qy   1024 TAACTATTGATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAA 1083
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1405 TAACTATTGATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAA 1464

Qy   1084 TGAGCATTCCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGA 1143
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1465 TGAGCATTCCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGA 1524

Qy   1144 ACACTAAAAGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAAGATGTC 1198
          ||||||||||||||||||||||||||||||||||||||||||||||||||||  |||
Db   1525 ACACTAAAAGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAAGTAGTC 1579
```

RESULT 7
ABK88015
ID ABK88015 standard; DNA; 1182 BP.
XX
AC ABK88015;
XX
DT 07-OCT-2002 (first entry)
XX
DE DNA encoding human alpha-1-antitrypsin (AAT) protein.
XX
KW Alpha-1-antitrypsin; AAT; human; gene; ds; protease inhibitor; malaria;
KW emphysema; asthma; chronic obstructive pulmonary disease; eczema;
KW cystic fibrosis; otitis media; otitis external; HIV; psoriasis;
KW human immunodeficiency virus; atopic dermatitis; muscular dystrophy;
KW herpes; ulceration; sepsis; rheumatoid arthritis; periodontal disease;
KW tumour metastasis; tumour angiogenesis; osteoporosis; Paget's disease;
KW glomerulonephritis; scleroderma; Alzheimer's disease; hypertension.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1182

```
FT                   /*tag= a
FT                   /product= "Alpha-1-antitrypsin"
FT                   /partial
FT                   /note= "No start or stop codon shown"
XX
PN   WO200250287-A2.
XX
PD   27-JUN-2002.
XX
PF   18-DEC-2001; 2001WO-US49256.
XX
PR   18-DEC-2000; 2000US-256699P.
PR   20-NOV-2001; 2001US-331966P.
XX
PA   (ARRI-) ARRIVA PHARM INC.
XX
PI   Barr PJ,  Gibson HL,  Pemberton P;
XX
DR   WPI; 2002-500631/53.
DR   P-PSDB; AAU99873.
XX
PT   Novel fusion protein useful for inhibiting protease activity associated
PT   with a disorder such as emphysema, asthma, comprises a first protease
PT   inhibitor comprising alpha 1-antitrypsin and a second protease
PT   inhibitor -
XX
PS   Disclosure; Page 24-25; 134pp; English.
XX
CC   This invention relates to a novel fusion protein comprising a first
CC   protease inhibitor comprising an alphal-antitrypsin or its functionally
CC   active portion and a second protease inhibitor or its functionally
CC   active protein. The fusion proteins of the invention may act as an
CC   inhibitor of protease activity. The fusion protein of the invention
CC   is useful for inhibiting protease activity associated with a disorder
CC   such as emphysema, asthma, chronic obstructive pulmonary disease,
CC   cystic fibrosis, otitis media, otitis external or HIV infection, or
CC   for treating an individual suffering from or at risk for a disease or
CC   disorder involving unwanted protease activity. The proteins are useful
CC   for treating dermatological diseases such as atopic dermatitis, eczema
CC   and psoriasis, in inflammatory responses to viral infection, and for
CC   treating herpes infection, corneal or epidermal ulceration, chronic
CC   non-healing wounds, sepsis, rheumatoid arthritis, periodontal disease,
CC   tumour metastasis and tumour angiogenesis, gastric ulceration,
CC   osteoporosis, Paget's disease, glomerulonephritis, scleroderma, malaria,
CC   bacterial infection, Alzheimer's disease, hypertension and muscular
CC   dystrophy. The present sequence represents the DNA encoding the human
CC   alpha-1-antitrypsin (AAT) protein used to create the fusion protein
CC   of the invention.
XX
SQ   Sequence 1182 BP; 369 A; 214 C; 229 G; 370 T; 0 other;

  Query Match             77.5%;  Score 1182;  DB 24;  Length 1182;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-287;
  Matches 1182;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 60

Qy      72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 120

Qy     132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 180

Qy     192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 240

Qy     252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 300

Qy     312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 360

Qy     372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 420

Qy     432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 480

Qy     492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 540

Qy     552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 600

Qy     612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 660

Qy     672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 720

Qy     732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTTACCAGACGAAGGTAAGCTT 791
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTTACCAGACGAAGGTAAGCTT 780

Qy     792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 840

Qy     852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 900

Qy     912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 960

Qy     972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1020

Qy    1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1080

Qy    1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1140

Qy    1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAAG 1193
           |||||||||||||||||||||||||||||||||||||||||||
Db    1141 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAAG 1182


RESULT 8
AAV41730
ID   AAV41730 standard; DNA; 1260 BP.
XX
AC   AAV41730;
XX
DT   20-NOV-1998  (first entry)
XX
DE   Codon-optimised RAmy3D signal fused to DNA encoding mature AAT.
XX
```

```
KW   Protein expression; monocotyledon plant cell;
KW   glycosylated alpha 1-antitrypsin; AAT; glycosylated antithrombin III;
KW   ATIII; human serum albumin; HSA; subtilisin BPN'; treatment; emphysema;
KW   antithrombotic; blood replacement; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1..75
FT                   /*tag= a
FT                   /note= "codon-optimised RAmy3D signal sequence"
FT   misc_feature    76..1260
FT                   /*tag= b
FT                   /note= "encodes mature AAT"
XX
PN   WO9836085-A1.
XX
PD   20-AUG-1998.
XX
PF   13-FEB-1998;   98WO-US03068.
XX
PR   13-FEB-1997;   97US-0038170.
PR   13-FEB-1997;   97US-0037991.
PR   13-FEB-1997;   97US-0038168.
PR   13-FEB-1997;   97US-0038169.
XX
PA   (PHYT-) APPLIED PHYTOLOGICS INC.
XX
PI   Rodriguez RL,  Sutliff TD;
XX
DR   WPI; 1998-467179/40.
XX
PT   Expressing mature, glycosylated proteins in monocotyledonous plant
PT   cells - from chimeric gene including signal peptide sequence,
PT   specifically therapeutic agents and industrial enzymes
XX
PS   Disclosure; Pages 34 iii-iv; 53pp; English.
XX
CC   The present sequence encodes a fusion protein of codon-optimised RAmy3D
CC   signal sequence/mature alpha1-antitrypsin (AAT). The protein is used
CC   to exemplify the invention. The specification describes a method for
CC   producing mature heterologous protein in monocotyledonous plant cells.
CC   The method comprises transforming the cells with a chimeric gene
CC   comprising a monocotyledon transcription regulator, inducible either
CC   during seed maturation or by adding/removing a small molecule, DNA
CC   encoding the heterologous protein, and DNA encoding a signal peptide,
CC   with the signal peptide causing secretion of the protein from the cell.
CC   Proteins expressed in this manner include mature glycosylated alpha
CC   1-antitrypsin (AAT) with a glycosylation pattern that significantly
CC   increases its serum half-life, mature glycosylated antithrombin III
CC   (ATIII), mature human serum albumin (HSA) having the native folding
CC   pattern as shown by bilirubin-binding characteristics, or mature active
CC   subtilisin BPN'. These proteins are useful therapeutically (e.g. AAT for
CC   treating emphysema, ATIII as antithrombotic and HSA as blood replacement)
CC   or as industrial enzymes (BPN' is used in detergents).
XX
SQ   Sequence 1260 BP; 287 A; 428 C; 350 G; 195 T; 0 other;

  Query Match             41.2%;  Score 628.4;  DB 19;  Length 1260;
  Best Local Similarity   70.7%;  Pred. No. 5.4e-148;
  Matches  836;  Conservative    0;  Mismatches  346;  Indels    0;  Gaps    0;

Qy      12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71
           || ||||| || |||||||||||| || || |||||||||||| || |||||||| |||||
Db      76 GAGGACCCGCAGGGCGACGCCGCCCAGAAGACCGACACCAGCCACCACGACCAGGACCAC 135

Qy      72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131
           ||||| || || || || || || ||||| |||||||| || ||     ||||  | ||
Db     136 CCGACGTTCAACAAGATCACCCCGAATTTGGCCGAATTCGCCTTCAGCCTGTACCGCCAG 195

Qy     132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191
             | || || ||    || || || ||||| || || || || ||    || ||||| ||

Db     196 CTCGCGCACCAGTCCAACTCCACCAACATCTTCTTCAGCCCGGTGAGCATCGCCACCGCC 255

Qy     192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
           ||||||||| ||    | ||||| || || || ||||| |||||||| | |||||  |
Db     256 TTCGCCATGCTGTCCCTGGGTACCAAGGCGGACACCCACGACGAGATCCTCGAAGGGCTG 315

Qy     252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311
           ||||| ||  |||| || ||||| || || || || |||||||| || || ||| || |
Db     316 AACTTCAACCTGACGGAGATCCCGGAGGCGCAGATCCACGAGGGCTTCCAGGAGCTGCTC 375

Qy     312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371
           || ||  | || || || || || || ||  | ||  | || || || |||||  | ||  ||
Db     376 AGGACGCTCAACCAGCCGGACTCCCAGCTCCAGCTCACCACCGGCAACGGGCTCTTCCTG 435

Qy     372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431
           || || ||  | ||  | || || || ||||| || ||||| ||||| || || ||
Db     436 TCCGAGGGCCTCAAGCTCGTCGATAAGTTCCTGGAGGACGTGAAGAAGCTCTACCACTCC 495

Qy     432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491
           ||||| || ||||| || || || || || ||||| || || ||||| || || || ||
Db     496 GAGGCGTTCACCGTCAACTTCGGGGACACCGAGGAGGCCAAGAAGCAGATCAACGACTAC 555

Qy     492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551
           || ||||| || |||||||| |||||||| ||||| || || ||||| ||  | || |||
Db     556 GTCGAGAAGGGGACCCAGGGCAAGATCGTGGACCTGGTCAAGGAATTGGACAGGGACACC 615

Qy     552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611
           |||||||| || || ||||| || || |||||||| |||||||| || || |||||||||
Db     616 GTCTTCGCGCTCGTCAACTACATCTTCTTCAAGGGCAAGTGGGAGCGCCCGTTCGAGGTG 675

Qy     612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671
           || || || || ||||| || || || || || || || || || ||||| || || |||
Db     676 AAGGACACCGAGGAGGAGGACTTCCACGTCGACCAGGTCACCACCGTCAAGGTCCCGATG 735

Qy     672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731
           ||||| || || || |||||||| || || || ||||| ||  |        |||||  |
Db     736 ATGAAGAGGCTCGGCATGTTCAACATCCAGCACTGCAAGAAGCTCTCCAGCTGGGTGCTC 795

Qy     732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTAAGCTT 791
           | ||||||||  | || ||||| || || || || ||  | || ||||| || |||||
Db     796 CTCATGAAGTACCTGGGGAACGCCACCGCCATCTTCTTCCTGCCGGACGAGGGCAAGCTC 855

Qy     792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851
           || ||  | ||||| ||| |||| || ||||| || || || ||  | ||||||||||||
Db     856 CAGCACCTGGAGAACGAGCTGACGCACGACATCATCACGAAGTTCCTGGAGAACGAGGAC 915

Qy     852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911
             | ||   ||||   || ||||| || ||| | || |||||||| || |||||| | ||
Db     916 AGGCGCTCCGCTAGCCTCCACCTCCCGAAGCTGAGCATCACCGGCACGTACGACCTGAAG 975

Qy     912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971
             ||  | |||||| | || || || || || ||    ||||| || ||  |    ||
Db     976 AGCGTGCTGGGCCAGCTGGGCATCACGAAGGTCTTCAGCAACGGCGCGGACCTCTCCGGC 1035

Qy     972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031
           || || || || || ||  | ||  |    || || || ||||| || ||  | || ||
Db    1036 GTGACGGAGGAGGCCCCCCTGAAGCTCTCCAAGGCCGTGCACAAGGCGGTGCTCACGATC 1095

Qy    1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091
           || || ||||| || || || ||||| || ||||||||||| || || || |||  |||
Db    1096 GACGAGAAGGGGACGGAAGCTGCCGGGGCCATGTTCCTGGAGGCCATCCCCATGTCCATC 1155

Qy    1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151
           || || || || || || || ||||| || ||||||||||||||||||||||||||| ||
Db    1156 CCGCCCGAGGTCAAGTTCAACAAGCCCTTCGTCTTCCTGATGATCGAGCAGAACACGAAG 1215

Qy    1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAAG 1193
           |||||  | || ||||| ||||| |||||||| || ||||||
Db    1216 AGCCCCCTCTTCATGGGGAAGGTCGTCAACCCCACGCAGAAG 1257

RESULT 9
AAQ89254
```

```
ID   AAQ89254 standard; cDNA; 1312 BP.
XX
AC   AAQ89254;
XX
DT   18-OCT-1995  (first entry)
XX
DE   Human alpha-1-trypsin cDNA.
XX
KW   Alpha-1-trypsin; protease-inhibitor; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             28..1258
FT                   /*tag= a
FT   sig_peptide     28..99
FT                   /*tag= b
FT   mat_peptide     100..1255
FT                   /*tag= c
XX
PN   US5399684-A.
XX
PD   21-MAR-1995.
XX
PF   20-MAY-1982;   82US-0380310.
XX
PR   20-MAY-1982;   82US-0380310.
PR   07-FEB-1984;   84US-0638980.
PR   03-MAR-1987;   87US-0022543.
PR   15-DEC-1987;   87US-0133190.
PR   16-SEP-1988;   88US-0246912.
PR   22-AUG-1989;   89US-0398288.
PR   11-MAR-1991;   91US-0666450.
PR   18-NOV-1992;   92US-0979556.
PR   02-JUL-1993;   93US-0086442.
XX
PA   (WASH-) WASHINGTON RES FOUND.
XX
PI   Davie EW,  Kurachi K,  Thirumalachary C,  Woo SLC;
XX
DR   WPI; 1995-130740/17.
DR   P-PSDB; AAR71969.
XX
PT   Human alpha1-antitrypsin (a1-AT) cDNA sequence - can be used for
PT   the expression of a1-AT
XX
PS   Claim 1; Fig.1; 15pp; English.
XX
CC   The sequence of a human alpha-1-antitrypsin cDNA clone is given in
CC   AAQ89254.  Expression of the cDNA in host cell transformants has
CC   allowed production of recombinant alpha-1-antitrypsin.
XX
SQ   Sequence 1312 BP; 339 A; 368 C; 324 G; 281 T; 0 other;

  Query Match            28.3%;  Score 432.2;  DB 16;  Length 1312;
  Best Local Similarity  60.4%;  Pred. No. 1.1e-98;
  Matches  713;  Conservative    0;  Mismatches  468;  Indels    0;  Gaps    0;

Qy      12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71
           || || || || || || || || || || || || ||    || || || || || ||
Db     100 GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 159

Qy      72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131
           || || || || || || || || || || || || || || ||      | ||  | ||
Db     160 CCAACCTTCAACAAGATCACCCCCAACTTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 219

Qy     132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191
           | || || ||    ||    || || || || ||    || ||    || || || ||
Db     220 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 279

Qy     192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
           || || ||| |      | || || || || || || || || || ||  | || ||  |
Db     280 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 339

Qy     252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311
           || || ||  | || || || || || || || || || || || || ||  | |
Db     340 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC 399

Qy     312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371
           | ||  | || || || ||     || | ||  | || || || || || ||  | ||  |
Db     400 CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTC 459

Qy     372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431
           || ||  | ||  | || || || ||  | || || || || ||  | || ||
Db     460 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 519

Qy     432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491
           || || || || || || || || || || || || || || || || || || |||||
Db     520 GAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 579

Qy     492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551
           || ||||| || || || || || || || || ||  | || || ||  | ||  | || ||
Db     580 GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 639

Qy     552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611
           || || || || || || || || || || || || || || ||||  | || || || ||
Db     640 GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC 699

Qy     612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671
           || || || || || || || || || || || || || || || || || || || |||
Db     700 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG 759

Qy     672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731
           |||||  |  | || ||||| || || || ||||| || ||  |       |||||  |
Db     760 ATGAAGCGTTTAGGCATGTTTAACATCCAGCATTGTAAGAAGCTGTCCAGCTGGGTGCTG 819

Qy     732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTAAGCTT 791
           | ||||| ||  | || || || || || || || ||  | || || || || || ||
Db     820 CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTA 879

Qy     792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851
           || ||  | || |||||  | || || || || || || ||  | || || || ||
Db     880 CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC 939

Qy     852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911
           | |    ||      | ||  | || ||  |    || || || || || ||  | ||
Db     940 AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG 999

Qy     912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971
           || |||| ||  | || || || || || ||    || || || ||  |    ||
Db    1000 AGCGTCCTAGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 1059

Qy     972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031
           || || || || || ||  | ||  |    || || || || || || ||  | || ||
Db    1060 GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC 1119

Qy    1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091
           || || || || || || || || || || || |||||  | || || || || |||   ||
Db    1120 GACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATC 1179

Qy    1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151
           |  || || || || || || ||||| || || ||  | ||||| || || || || || ||
Db    1180 CGCCCCGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAG 1239

Qy    1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAA 1192
           ||  | || ||||| || || || || || || || || ||
Db    1240 TCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA 1280


RESULT 10
AAV28471
ID   AAV28471 standard; cDNA; 1312 BP.
XX
AC   AAV28471;
XX
DT   21-AUG-1998  (first entry)
```

```
XX
DE   Nucleotide sequence of the alpha-1-antitrypsin.
XX
KW   Human alpha-1-antitrypsin; ATR-1; antibody; ATR-1 deficiency; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             28..1257
FT                   /*tag= a
FT                   /product= "alpha-1-antitrypsin"
XX
PN   US5736379-A.
XX
PD   07-APR-1998.
XX
PF   07-JUN-1995;   95US-0479545.
XX
PR   20-MAY-1982;   82US-0380310.
PR   07-FEB-1984;   84US-0638980.
PR   03-MAR-1987;   87US-0022543.
PR   15-DEC-1987;   87US-0133190.
PR   16-SEP-1988;   88US-0246912.
PR   22-AUG-1989;   89US-0398288.
PR   11-MAR-1991;   91US-0666450.
PR   18-NOV-1992;   92US-0979556.
PR   02-JUL-1993;   93US-0086442.
PR   12-DEC-1994;   94US-0361689.
XX
PA   (WASH-) WASHINGTON RES FOUND.
XX
PI   Davie EW,  Kurachi K,  Thirumalachary C,  Woo SLC;
XX
DR   WPI; 1998-239214/21.
DR   P-PSDB; AAW56709.
XX
PT   DNA encoding alpha-1 anti-trypsin - useful for, e.g. producing
PT   recombinant alpha-1 anti-trypsin
XX
PS   Claim 1; Fig 1; 15pp; English.
XX
CC   This is the nucleotide sequence encoding the novel human
CC   alpha-1-antitrypsin (ATR-1) protein.  Its products are useful for
CC   producing recombinant ATR-1 polypeptides, which can be used to prepare
CC   antibodies for detecting ATR-1 variants in the blood, as ligands in
CC   assays for ATR-1, and to treat ATR-1 deficiency.
XX
SQ   Sequence 1312 BP; 339 A; 368 C; 324 G; 281 T; 0 other;

  Query Match             28.3%;  Score 432.2;  DB 19;  Length 1312;
  Best Local Similarity   60.4%;  Pred. No. 1.1e-98;
  Matches  713;  Conservative    0;  Mismatches  468;  Indels    0;  Gaps    0;

Qy      12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71
           || || || || || || || || || || || || ||    || || || || || ||
Db     100 GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 159

Qy      72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131
           || || || || || || || || || || || || || || ||     | ||   | ||
Db     160 CCAACCTTCAACAAGATCACCCCCAACTTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 219

Qy     132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191
            | || || ||    ||    || || || || ||    || ||    || || || ||
Db     220 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 279

Qy     192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
           || || ||| |     | || || || || || || || || ||   | || ||   |
Db     280 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 339

Qy     252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311
           || || ||  | || || || || || || || || || || || || || || ||  |  |
Db     340 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC 399

Qy     312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371
           | ||  | || || || ||    ||  | ||  | || || || || ||  | ||  |
Db     400 CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTC 459

Qy     372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431
             || ||  | ||  | || || || ||  | || || || || ||  | || ||
Db     460 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 519

Qy     432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491
           || || || || || || || || || || || || || || || || || || ||||||
Db     520 GAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 579

Qy     492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551
           || ||||| || || || || || || || || || || || || ||  | ||  | || ||
Db     580 GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 639

Qy     552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611
           || || || || || || || || || || || || || || || || ||||||  | || ||
Db     640 GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC 699

Qy     612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671
           || || || || || || || || || || || || || || || || || || || |||
Db     700 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG 759

Qy     672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731
           |||||  |  | || ||||| || || || ||||| || ||  |       |||||  |
Db     760 ATGAAGCGTTTAGGCATGTTTAACATCCAGCATTGTAAGAAGCTGTCCAGCTGGGTGCTG 819

Qy     732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTAAGCTT 791
           | ||||| ||  | || || || || || || || || ||  | || || || || || ||
Db     820 CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTA 879

Qy     792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851
           || ||  | || |||||  | || || || || || || || ||  | || || || ||
Db     880 CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC 939

Qy     852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911
           |  |    ||     | ||  | || ||  |    || || || || || ||  | ||
Db     940 AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG 999

Qy     912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971
             ||  |||| ||  | || || || || || || ||    || || || ||  |    ||
Db    1000 AGCGTCCTAGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 1059

Qy     972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031
           || || || || || ||  | ||  |    || || || || || || ||  | || ||
Db    1060 GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC 1119

Qy    1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091
           || || || || || || || || || || || |||||  | || || || || |||   ||
Db    1120 GACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATC 1179

Qy    1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151
           |  || || || || || || || |||||| || || ||  | |||||| || || || || ||
Db    1180 CGCCCCGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAG 1239

Qy    1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAA 1192
             ||  | || |||| || || || || || || || ||
Db    1240 TCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA 1280


RESULT 11
AAZ90199
ID   AAZ90199 standard; cDNA; 1312 BP.
XX
AC   AAZ90199;
XX
DT   19-MAY-2000  (first entry)
XX
DE   Human alpha1-antitrypsin nucleotide sequence.
XX
KW   Alpha1-antitrypsin; neutrophil elastase inhibitor; human; ss;
KW   chronic obstructive pulmonary emphysema; infantile liver cirrhosis.
```

```
XX
OS   Homo sapiens.
XX
PN   US6025161-A.
XX
PD   15-FEB-2000.
XX
PF   20-JAN-1998;   98US-0009581.
XX
PR   07-JUN-1995;   95US-0479545.
PR   20-MAY-1982;   82US-0380810.
PR   07-FEB-1984;   84US-0638980.
PR   03-MAR-1987;   87US-0022543.
PR   15-DEC-1987;   87US-0133190.
PR   16-SEP-1988;   88US-0246912.
PR   22-AUG-1989;   89US-0398288.
PR   11-MAR-1991;   91US-0666450.
PR   18-NOV-1992;   92US-0979556.
PR   02-JUL-1993;   93US-0086442.
XX
PA   (WASH-) WASHINGTON RES FOUND.
XX
PI   Woo SLC,  Thirumalachary C,  Kurachi K,  Davie EW;
XX
DR   WPI; 2000-181811/16.
DR   P-PSDB; AAY78890.
XX
PT   Preparing alpha1-antitrypsin for inhibiting neutrophil elastase
PT   involves transfecting host cell with vector comprising
PT   alpha1-antitrypsin DNA sequence that hybridizes to human
PT   alpha1-antitrypsin cDNA, or its complement -
XX
PS   Claim 1; Fig 1; 16pp; English.
XX
CC   This sequence represents the human alpha1-antitrypsin nucleotide
CC   sequence. Alpha1-antitrypsin is an important protease inhibitor, the
CC   major function of which is to inhibit neutrophil elastase. Low levels of
CC   alpha1-antitrypsin in the blood are associated with chronic obstructive
CC   pulmonary emphysema and infantile liver cirrhosis. A vector comprising a
CC   mammalian alpha1-antitrypsin DNA sequence that hybridises to human
CC   alpha1-antitrypsin cDNA can be introduced into a host cell in a method
CC   for the production of alpha1-antitrypsin.
XX
SQ   Sequence 1312 BP; 339 A; 368 C; 324 G; 281 T; 0 other;

  Query Match           28.3%;  Score 432.2;  DB 21;  Length 1312;
  Best Local Similarity  60.4%;  Pred. No. 1.1e-98;
  Matches  713;  Conservative    0;  Mismatches  468;  Indels    0;  Gaps    0;

Qy      12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71
           || || || || || || || || || || || || || ||    || || || || || ||
Db     100 GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 159

Qy      72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131
           || || || || || || || || || || || || || || || ||     | ||  | ||
Db     160 CCAACCTTCAACAAGATCACCCCCAACTTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 219

Qy     132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191
           | || || ||   ||    || || || || ||    || ||    || || || ||
Db     220 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 279

Qy     192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
           || || ||| |     | || || || || || || || || || ||  | || ||  |
Db     280 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 339

Qy     252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311
           || || ||  | || || || || || || || || || || || || || ||  |  |
Db     340 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC 399

Qy     312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371
           | ||  | || || || ||    ||  | ||  | || || || || ||  | ||  |
Db     400 CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTC 459

Qy     372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431
           || ||  | ||  | || || || ||  | || || || || ||  | || ||
Db     460 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 519

Qy     432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491
           || || || || || || || || || || || || || || || || || || || |||||
Db     520 GAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 579

Qy     492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551
           || ||||| || || || || || || || || ||  | || || ||  | ||  | || ||
Db     580 GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 639

Qy     552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611
           || || || || || || || || || || || || || || ||||||  | || || || ||
Db     640 GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC 699

Qy     612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671
           || || || || || || || || || || || || || || || || || || || |||
Db     700 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG 759

Qy     672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731
           |||||  |  | || |||||  || || || |||||  || ||  |        |||||  |
Db     760 ATGAAGCGTTTAGGCATGTTTAACATCCAGCATTGTAAGAAGCTGTCCAGCTGGGTGCTG 819

Qy     732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTAAGCTT 791
           | ||||| ||  | || || || || || || || ||  | || || || || || ||
Db     820 CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTA 879

Qy     792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851
           || ||  | || ||||  | || || || || || || || || || || || || || ||
Db     880 CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC 939

Qy     852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911
           |  |    ||     | ||  | || ||  |    || || || || || ||  | ||
Db     940 AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG 999

Qy     912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971
           ||  |||| ||  | || || || || || ||    || || || ||  |  ||
Db    1000 AGCGTCCTAGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 1059

Qy     972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031
           || || || || || ||  | ||  |    || || || || || || || ||  | || ||
Db    1060 GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC 1119

Qy    1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091
           || || || || || || || || || || || |||||  | || || || || |||  ||
Db    1120 GACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATC 1179

Qy    1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151
           |  || || || || || || ||||| || || ||  | ||||| || || || || ||
Db    1180 CGCCCCGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAG 1239

Qy    1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAA 1192
              ||  | || ||||| || || || || || || || ||
Db    1240 TCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA 1280


RESULT 12
AAS45052
ID   AAS45052 standard; cDNA; 1367 BP.
XX
AC   AAS45052;
XX
DT   18-DEC-2001  (first entry)
XX
DE   cDNA encoding novel human secretory protein, Seq ID No 133.
XX
KW   Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
KW   ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
KW   transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
KW   amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
KW   ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
KW   gut protection; lung; liver fibrosis; immune deficiency; infection;
```

KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
KW fertility; analgesic; pain; antigen; ss.
XX
OS Homo sapiens.
XX
PN WO200166689-A2.
XX
PD 13-SEP-2001.
XX
PF 05-MAR-2001; 2001WO-US04942.
XX
PR 07-MAR-2000; 2000US-0519705.
PR 19-MAY-2000; 2000US-0574454.
PR 17-JUN-2000; 2000US-0596193.
PR 14-JUL-2000; 2000US-0616847.
PR 19-SEP-2000; 2000US-0665363.
PR 20-OCT-2000; 2000US-0693267.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
XX
DR WPI; 2001-589934/66.
DR P-PSDB; AAU28152.
XX
PT Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis and treatment of
PT cancer, neurological, inflammatory, and autoimmune disorders -
XX
PS Claim 1; SEQ ID No 133; 107pp; English.
XX
CC The invention relates to novel isolated human secreted polypeptides (I)
CC and polynucleotides (II). (I) and (II) are useful for treating
CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC involved in increasing haematopoiesis, stem cell survival, bone growth
CC and remodeling. (I), (II) and modulators of (II) are useful for
CC prophylaxis or treatment of one or more cancers. (II) is also useful for
CC creating transgenic animals useful for studying the in vivo activities of
CC the polypeptide as well as for studying modulators of the polypeptides.
CC (I) induces the proliferation of neural cells and regeneration of nerve
CC and brain tissue and is useful for the treatment of central and
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis and is useful for treating myeloid
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC tissue growth, and in tissue repair, healing of burns, incisions,
CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
CC disorders, or periodontal disease. Furthermore, (I) is also useful for
CC gut protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, various immune deficiencies and
CC disorders including severe combined immunodeficiency (SCID), bacterial or
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC reactions and conditions, such as asthma or other respiratory problems.
CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
CC fertility, metabolism, catabolism, anabolism, storage or elimination of
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
CC analgesic effects or other pain reducing effects, immunoglobulin like
CC activity and can act as an antigen in a vaccine composition to raise an
CC immune response. AAS44920-AAS45295 represent novel human secreted protein
CC coding sequences of the invention.
XX
SQ Sequence 1367 BP; 357 A; 392 C; 323 G; 295 T; 0 other;

Query Match 28.1%; Score 429.2; DB 22; Length 1367;
Best Local Similarity 60.7%; Pred. No. 6.6e-98;
Matches 718; Conservative 0; Mismatches 463; Indels 1; Gaps 1;

```
Qy    12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71
Db   105 GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 164

Qy    72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131
Db   165 CCAACCTTCAACAAGATCACCCCCAACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 224

Qy   132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191
Db   225 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 284

Qy   192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
Db   285 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 344

Qy   252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311
Db   345 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC 404

Qy   312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371
Db   405 CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTC 464

Qy   372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431
Db   465 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 524

Qy   432 GAGGCTTTTACCGTTAATTTTGGTGAT-ACTGAGGAAGCTAAAAAGCAAATTAATGATTA 490
Db   525 GAAGCCTTCACTGTCAACTTCGGGGATCACCGAAGAGGCCAAGAAACAGATCAACGATTA 584

Qy   491 TGTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATAC 550
Db   585 CGTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACAC 644

Qy   551 CGTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGT 610
Db   645 AGTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCTTTTGAAGT 704

Qy   611 TAAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAAT 670
Db   705 CAAGGACACCGAGGACGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTCCCTAT 764

Qy   671 GATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTT 730
Db   765 GATGAAGCGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTACT 824

Qy   731 ATTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTAAGCT 790
Db   825 GCTAATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTACCTGATGAGGGGAAACT 884

Qy   791 TCAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGA 850
Db   885 ACAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGA 944

Qy   851 TCGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAA 910
Db   945 CAGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAA 1004

Qy   911 ATCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGG 970
Db  1005 GAGCGTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGG 1064

Qy   971 TGTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTAT 1030
Db  1065 GGTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCAT 1124

Qy  1031 TGATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCAT 1090
Db  1125 CGACGAGAAGGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCAATGTCTAT 1184

Qy  1091 TCCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAA 1150
```

```
Db   1185 CCCCCCAGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAA 1244

Qy   1151 AAGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAA 1192

Db   1245 GTCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA 1286
```

RESULT 13
AAV41726
ID AAV41726 standard; DNA; 1185 BP.
XX
AC AAV41726;
XX
DT 20-NOV-1998 (first entry)
XX
DE Native coding sequence of mature alpha1-antitrypsin (AAT).
XX
KW Protein expression; monocotyledon plant cell;
KW glycosylated alpha 1-antitrypsin; AAT; glycosylated antithrombin III;
KW ATIII; human serum albumin; HSA; subtilisin BPN'; treatment; emphysema;
KW antithrombotic; blood replacement; ss.
XX
OS Homo sapiens.
XX
PN WO9836085-A1.
XX
PD 20-AUG-1998.
XX
PF 13-FEB-1998; 98WO-US03068.
XX
PR 13-FEB-1997; 97US-0038170.
PR 13-FEB-1997; 97US-0037991.
PR 13-FEB-1997; 97US-0038168.
PR 13-FEB-1997; 97US-0038169.
XX
PA (PHYT-) APPLIED PHYTOLOGICS INC.
XX
PI Rodriguez RL, Sutliff TD;
XX
DR WPI; 1998-467179/40.
DR P-PSDB; AAW59839.
XX
PT Expressing mature, glycosylated proteins in monocotyledonous plant
PT cells - from chimeric gene including signal peptide sequence,
PT specifically therapeutic agents and industrial enzymes
XX
PS Disclosure; Page 29; 53pp; English.
XX
CC The present sequence represents the native coding sequence of mature
CC alpha1-antitrypsin (AAT). The protein is used to exemplify the
CC invention. The specification describes a method for producing mature
CC heterologous protein in monocotyledonous plant cells. The method
CC comprises transforming the cells with a chimeric gene comprising a
CC monocotyledon transcription regulator, inducible either during seed
CC maturation or by adding/removing a small molecule, DNA encoding the
CC heterologous protein, and DNA encoding a signal peptide, with
CC the signal peptide causing secretion of the protein from the cell.
CC Proteins expressed in this manner include mature glycosylated alpha
CC 1-antitrypsin (AAT) with a glycosylation pattern that significantly
CC increases its serum half-life, mature glycosylated antithrombin III
CC (ATIII), mature human serum albumin (HSA) having the native folding
CC pattern as shown by bilirubin-binding characteristics, or mature active
CC subtilisin BPN'. These proteins are useful therapeutically (e.g. AAT for
CC treating emphysema, ATIII as antithrombotic and HSA as blood replacement)
CC or as industrial enzymes (BPN' is used in detergents).
XX
SQ Sequence 1185 BP; 328 A; 324 C; 283 G; 250 T; 0 other;

Query Match 28.1%; Score 429; DB 19; Length 1185;
Best Local Similarity 60.2%; Pred. No. 7.1e-98;
Matches 711; Conservative 0; Mismatches 470; Indels 0; Gaps 0;

```
Qy     12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71
Db      1 GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 60

Qy     72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131
Db     61 CCAACCTTCAACAAGATCACCCCCAACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 120

Qy    132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191
Db    121 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 180

Qy    192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
Db    181 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 240

Qy    252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311
Db    241 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC 300

Qy    312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371
Db    301 CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTC 360

Qy    372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431
Db    361 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 420

Qy    432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491
Db    421 GAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 480

Qy    492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551
Db    481 GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 540

Qy    552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611
Db    541 GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC 600

Qy    612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671
Db    601 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG 660

Qy    672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731
Db    661 ATGAAGCGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTG 720

Qy    732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTAAGCTT 791
Db    721 CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTA 780

Qy    792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851
Db    781 CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC 840

Qy    852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911
Db    841 AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG 900

Qy    912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971
Db    901 AGCGTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 960

Qy    972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031
Db    961 GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC 1020

Qy   1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091
Db   1021 GACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATC 1080

Qy   1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151
```

```
Db   1081 CCCCCCGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAG 1140

Qy   1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAA 1192

Db   1141 TCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA 1181
```

RESULT 14
AAQ31403
ID AAQ31403 standard; DNA; 1352 BP.
XX
AC AAQ31403;
XX
DT 23-MAR-1993 (first entry)
XX
DE Human alpha-1 antitrypsin.
XX
KW Plasmid; pCMV4; liposome; antiprotease; lung; emphysema;
KW adult respiratory distress syndrome; ARDS; ss.
XX
OS Homo sapiens.
XX
PN WO9219730-A.
XX
PD 12-NOV-1992.
XX
PF 27-MAR-1992; 92WO-US02465.
XX
PR 24-APR-1991; 91US-0690283.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
PI Brigham K, Canonico A, Conary J, Meyrick B;
XX
DR WPI; 1992-398857/48.
XX
PT Human alpha-1 anti-trypsin contg. plasmid - for treatment of
PT anti-protease deficiency in emphysema and other lung diseases
XX
PS Disclosure; Fig 6a-6b; 32pp; English.
XX
CC A plasmid consisting of a pCMV4 expression vector including a coding
CC sequence of human alpha-1 antitrypsin may be incorporated into
CC liposomes capable of targeting specific tissue. The plasmid is then
CC capable of expression of the gene extrachromosomally in the cells of
CC the target tissue and is unincorporable into the chromosome of the
CC cells of the target tissue. Thus, the liposome including the
CC plasmid can be used in a method for treating a deficiency of the
CC gene product in cells of the target tissue.
CC The specific use of the human alpha-1 antitrypsin is significant as
CC this antiprotease is important in protecting the lungs against
CC emphysema. The adult respiratory distress syndrome (ARDS) is thought
CC to involve a relative deficiency of antiprotease activity.
CC Therefore, the delivery of a functioning alpha-1 antiprotease
CC gene to the lungs can be therapeutic in many human conditions
CC characterised by injury of the lungs.
XX
SQ Sequence 1352 BP; 349 A; 386 C; 325 G; 292 T; 0 other;

Query Match 28.1%; Score 429; DB 13; Length 1352;
Best Local Similarity 60.2%; Pred. No. 7.4e-98;
Matches 711; Conservative 0; Mismatches 470; Indels 0; Gaps 0;

```
Qy     12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71

Db     92 GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 151

Qy     72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131

Db    152 CCAACCTTCAACAAGATCACCCCCAACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 211

Qy    132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191

Db    212 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 271

Qy    192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251

Db    272 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 331

Qy    252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311

Db    332 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC 391

Qy    312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371

Db    392 CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTC 451

Qy    372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431

Db    452 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 511

Qy    432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491

Db    512 GAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 571

Qy    492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551

Db    572 GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 631

Qy    552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611

Db    632 GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC 691

Qy    612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671

Db    692 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG 751

Qy    672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731

Db    752 ATGAAGCGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTG 811

Qy    732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTAAGCTT 791

Db    812 CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTA 871

Qy    792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851

Db    872 CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC 931

Qy    852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911

Db    932 AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG 991

Qy    912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971

Db    992 AGCGTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 1051

Qy    972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031

Db   1052 GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC 1111

Qy   1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091

Db   1112 GACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATC 1171

Qy   1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151

Db   1172 CCCCCCGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAG 1231

Qy   1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAA 1192

Db   1232 TCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA 1272
```

RESULT 15

ABL67511
ID ABL67511 standard; DNA; 1352 BP.
XX
AC ABL67511;
XX
DT 15-MAY-2002 (first entry)
XX
DE Thyroid cancer related gene sequence SEQ ID NO:5848.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
PS Claim 1; SEQ ID 5848; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 1352 BP; 349 A; 386 C; 325 G; 292 T; 0 other;

Query Match 28.1%; Score 429; DB 24; Length 1352;
Best Local Similarity 60.2%; Pred. No. 7.4e-98;
Matches 711; Conservative 0; Mismatches 470; Indels 0; Gaps 0;

```
Qy     12 GAAGACCCTCAAGGCGACGCCGCTCAAAAAACCGACACCAGTCATCACGACCAAGACCAT 71
          || || || || || || || || || || || || || ||    || || || || || ||
Db     92 GAGGATCCCCAGGGAGATGCTGCCCAGAAGACAGATACATCCCACCATGATCAGGATCAC 151

Qy     72 CCGACTTTTAATAAAATTACTCCAAATTTAGCCGAATTTGCTTTTTCTTTGTATAGACAA 131
          || || || || || || || || || || | || || || || ||     | ||  | ||
Db    152 CCAACCTTCAACAAGATCACCCCCAACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 211

Qy    132 TTAGCTCATCAAAGTAATTCTACTAACATTTTTTTTAGTCCTGTTTCTATTGCCACTGCT 191
           | || || ||    ||    || || || || ||    || ||    || || || ||
Db    212 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCTCCCCAGTGAGCATCGCTACAGCC 271

Qy    192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
          || || ||| |     | || || || || || || || || || || ||  | || ||  |
Db    272 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG 331

Qy    252 AACTTTAATTTGACCGAAATCCCAGAAGCCCAAATTCACGAGGGTTTTCAAGAGTTGTTG 311
          || || ||  | || || || || || || || || || || || || || ||  |  |
Db    332 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTC 391

Qy    312 AGAACTTTGAATCAACCTGATTCTCAATTGCAATTAACTACTGGTAACGGTTTATTTTTG 371
           | ||  | || || || ||    ||  | ||  | || || || || ||  | ||  |
Db    392 CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCCTGTTCCTC 451

Qy    372 TCTGAAGGTTTAAAATTGGTTGACAAATTCCTAGAAGACGTCAAGAAACTATATCATAGT 431
             || ||  | ||  | || || || ||  | || || || || ||  | || ||
Db    452 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCA 511

Qy    432 GAGGCTTTTACCGTTAATTTTGGTGATACTGAGGAAGCTAAAAAGCAAATTAATGATTAT 491
          || || || || || || || || || || || || || || || || || || |||||
Db    512 GAAGCCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 571

Qy    492 GTTGAGAAAGGCACCCAGGGTAAGATCGTTGACCTAGTTAAAGAATTAGATCGTGATACC 551
          || ||||| || || || || || || || || ||  | || || ||  | ||  | || ||
Db    572 GTGGAGAAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACA 631

Qy    552 GTCTTCGCACTAGTTAACTATATTTTTTTCAAGGGTAAGTGGGAACGTCCTTTCGAGGTT 611
          || || || || || || || || || || || || || || ||||  | || || || ||
Db    632 GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC 691

Qy    612 AAAGATACTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAATG 671
          || || || || || || || || || || || || || || || || || || || |||
Db    692 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG 751
```

```
Qy    672 ATGAAAAGACTGGGTATGTTCAATATTCAACATTGCAAAAAATTAAGTTCTTGGGTCTTA 731
          |||||  |  | ||  ||||| || || || || || || ||  |       |||||  |
Db    752 ATGAAGCGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTG 811

Qy    732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTTTTTTTACCAGACGAAGGTAAGCTT 791
           | ||||| ||  | || || || || || || || || ||  | || || || || || || ||
Db    812 CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTA 871

Qy    792 CAACATTTAGAGAATGAGTTGACTCATGACATTATTACTAAATTTTTAGAGAACGAGGAT 851
          || ||  | || |||||  | || || || || || || || || ||  | || || || ||
Db    872 CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTTCCTGGAAAATGAAGAC 931

Qy    852 CGTCGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911
           |  |    ||     | ||  | || ||  |    || || || || || ||  | ||
Db    932 AGAAGGTCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG 991

Qy    912 TCTGTTTTAGGCCAGTTAGGTATTACCAAAGTTTTTTCTAACGGTGCCGATTTGAGTGGT 971
             ||  | || ||  | || || || || || ||    || || || ||  |    ||
Db    992 AGCGTCCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 1051

Qy    972 GTTACTGAAGAAGCTCCATTAAAATTGAGTAAAGCTGTTCACAAAGCCGTCTTAACTATT 1031
          || || || || || ||  | ||  |    || || || || || || ||  | || ||
Db   1052 GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC 1111

Qy   1032 GATGAAAAGGGTACCGAGGCCGCCGGCGCTATGTTCCTGGAAGCTATTCCAATGAGCATT 1091
          || || || || || || || || || || ||||| | || || || || ||| ||
Db   1112 GACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGTCTATC 1171

Qy   1092 CCACCAGAAGTTAAATTTAATAAACCATTCGTTTTTCTGATGATCGAGCAGAACACTAAA 1151
          || || || || || || || ||||| || || ||  | ||||| || || || || ||
Db   1172 CCCCCCGAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAG 1231

Qy   1152 AGCCCATTGTTTATGGGTAAGGTTGTCAACCCAACTCAGAA 1192
             ||  | || ||||| || || || || || || || ||
Db   1232 TCTCCCCTCTTCATGGGAAAAGTGGTGAATCCCACCCAAAA 1272
```

Search completed: December 6, 2002, 21:11:40
Job time : 239.5 secs